DOCUMENT RESUME

ED 300 482 UD 026 464

AUTHOR Archambault, Francis X., Jr.

TITLE Instructional Setting and Other Design Features of Compensatory Education Programs.

PUB DATE Apr 87

NOTE 110p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, DC, April 23, 1987).

PUB TYPE Speeches/Conference Papers (150) -- Information Analyses (070) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC05 Plus Postage.

DESCRIPTORS Ability Grouping; Basic Skills; *Classroom Environment; Class Size; *Compensatory Education; Elementary Education; Elementary School Students; Heterogeneous Grouping; Language Arts; Literature Reviews; Program Design; Program Effectiveness; Remedial Mathematics; *Remedial Programs; Remedial Reading; Small Classes; Specialists; *Staff Utilization; Teacher Aides

IDENTIFIERS Education Consolidation Improvement Act Chapter 1

ABSTRACT

A review of the literature on compensatory education programs was conducted to determine the effects of their following design features: (1) instructional setting; (2) class size; (3) ability grouping; and (4) various types of instructors. Remedial instruction in reading, language arts, and mathematics for elementary school students was emphasized. Findings and recommendations include the following: (1) instructional setting is not directly related to student outcomes, so local conditions should determine whether an in-class or pull-out approach is used; (2) while it is clear that teachers need help in meeting the needs of the disadvantaged, there does not seem to be an optimal combination of instructional specialists, aides, or other teachers, so assignments should be based on staff competencies and student needs; and (3) small classes of 15 or fewer heterogeneously grouped students are preferred. One table illustrating the results of the Various Perspectives on the Pull-out Findings of the Instructional Dimensions Study is included. A 16-page list of references is also included. (FMW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

11-1-88 UD

ED300482

Instructional Setting and Other Design Features of Compensatory Education Programs

Prepared by

Francis X. Archambault, Jr.
University of Connecticut

April 1987

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

F.X. Archambault, jr.
U Conn

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented for the Annual Meeting of the American Educational Research Association, Washington, DC, April 23, 1987.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

UD026464

BEST COPY AVAILABLE

The past decade has been a particularly fruitful one for research on effective educational programs. We have learned a great deal about factors which distinguish effective from not-so-effective schools. We have also uncovered a wealth of information about teaching practices which contribute to improved student performance. Much of what we have learned from the effective schools literature can be used in schools serving the academically disadvantaged. In fact, such resarch has usually been conducted in this type of environment. Fortunately, some of the effective teaching research is also transportable to these schools and the classrooms within them. Nonetheless, much remains to be learned about the types of classroom practices and strategies which will lead to improved performance for compensatory education students.

This paper is concerned with what the literature says about the effects of certain design features of compensatory education programs. Specifically, it focuses on the effect of instructional setting (i.e., whether instruction is delivered inclass or outside the regular classroom in a pullout setting), the size of the class and the types of groups in which instruction is offered, and the effectiveness of

various types of instructors. Because the large majority of the dollars allocated to compensatory education provides remedial instruction in reading, language arts, and mathematics for elementary school students, these content areas and grade levels are emphasized. Further, because most compensatory instruction is provided in a pullout setting, and since the choice of setting may influence other design features, this design feature is emphasized.

THE OVERALL EFFECT OF COMPENSATORY EDUCATION

In 1965 Congress passed the Elementary and Secondary Education Act (ESEA). Title I of this Act was designed to improve the educational opportunities of the poor and educationally disadvantaged by providing federal dollars to all states and, through them, to local school districts most in need. Given the proposed large investment in Title I and the concern of many Congressional leaders that federal funds would not be used as intended, ESEA required that the Title I program be evaluated. These evaluations were the first systematic studies of the overall effect of compensatory education programs. However, even before these studies had begun, the Coleman Report conclusion that "schools bring little

influence to bear on a child's achievement that is independent of his background and general social control" (Coleman et al., 1966, p. 274) was being interpreted by many as an indictment of Title I (Stickney & Marcus, 1985).

Wargo, Tallmadge, Michaels, Lipe, and Morris (1972) provide the best review of the early evaluations of Title I. They concluded that there was little evidence to support the contention that Title I had a positive impact on participating students. Stickney and Plunkett (1982) have argued that the disappointing findings of these early evaluations may have been different had the scores for all participants not been lumped together. That is, some children did appear to benefit from Title I, but, when their scores were combined with those of students who did not, the positive effect was neutralized. McLaughlin (1975) reached a different conclusion. She argued that the initial attempts to evaluate Title I were unsuccessful because of the newness of the program, the diverse nature of the instructional components, and problems in obtaining cooperation between the federally sponsored evaluators and local school personnel.

Carter (1984) has reported on a more recent study

of Title I, one which he claims is the largest study of elementary education ever conducted. This research, known as the Sustaining Effects Study, focused on about 1,100 to 3,000 Title I students, about 1,300 to 2,000 needy students, and about 1,700 to 2,500 regular students in each of the grades 1 through 6. Title I students were reported by their school to be receiving some Title I services, needy students were reported by their teachers as being in need of Title I but not receiving them, and regular students were reported as not needing Title I services and not receiving them. The major results of this investigation were the following:

(1) Compared to needy students, Title I students showed significant gains in mathematics for each of the six grade levels.

(2) Compared to needy students, Title I students showed significant gains in reading for grades 1 through 3 but not for grades 4 through 6.

(3) The largest gains in reading and mathematics occurred in the first grade.

(4) The rate of gain for Title I students was at least equal to that for regular students,

whereas the rate of gain for needy students who did not receive Title I was not as great as that of either Title I students or regular students.

(5) Title I services were found to be effective for students who were only moderately disadvantaged (near average in achievement) but they did not improve the relative achievement of the most disadvantaged part of the school population.

(6) By the time students reached junior high school, there was no evidence of sustained or delayed effects of Title I.

How should these results be interpreted? Does compensatory education have a sufficiently positive impact on students to justify the enormous amount of money (over $40 billion) that has been allocated to it?

Wang (1980) reported that although compensatory education has had a positive impact on achievement growth, the effects of compensatory education efforts are not large. In fact, she concluded that the "evidence for positive effects of special instruction on achievement growth is sparse" (p. 8). Hallinan (1986) concludes that the results of the Sustaining

Effects Study are disappointing. She argues that "the desired outcome of a program of this magnitude would be steady, significant achievement gains that are sustained throughout a student's school career" (p. 4). Cooley (1981) concludes that "on balance, Title I is not producing the kind of achievement impact that people had hoped it would" (p. 300).

Stickney and Plunkett (1982) offer a different conclusion for the Sustaining Effects findings which they contend is influenced by data from the National Assessment of Educational Progress (NAEP). This latter research (NAEP, 1981) found that between 1970 and 1980 black nine-year-olds gained 9.9 percentage points on their reading assessment measure whereas whites improved their achievement by only 2.8 percentage points. NAEP also reported significantly greater gains between 1970 and 1980 for students in Title I eligible schools for all three grade levels tested. According to NAEP "these significant changes and the overall pattern of a narrowing gap for most population groups at all ages strongly suggest that students in Title I schools are improving at a faster rate than students in non-Title I schools" (cited in Stickney and Plunkett, 1982, p. 381). Stickney and Plunkett caution that since the NAEP conclusions are

based on correlations, one cannot argue that Title I has caused greater pupil achievement. However, they do conclude that the Sustaining Effects and NAEP studies "have finally reported what a host of smaller scale evaluations have suggested for years: Title I is associated with greater than expected pupil gains" (p. 380). Forbes (1985) is even more sanguine about the positive impact of compensatory education programs. In a relatively recent Phi Delta Kappa article Forbes concluded the following:

> Many things that happened in the Seventies could have dramatically affected student achievement in reading, among them: Title I, desegration, sensitivity to serving students who had been underserved in the past, movements to improve public education in the southeastern states, and a focus on social concerns. Of these, any could have had a major impact on student performance. But the movement that stands out above all others during the Seventies is the effort to better educate the economically disadvantaged. Title I and all the associated national and state efforts probably played a central role in the improvement of reading skills among those students who historically performed less well. (1985, p. 544)

In 1981 the Education and Improvement Act (ECIA) was enacted, thereby replacing Title I with Chapter I. Although we have little systematic data on the effect of this relatively new Act, Carter (1984) claims that the results "for Title I will apply equally well to the new Chapter I compensatory education program" (p. 4). Cooley (1981), however, contends that the search for overall effects of Title I (now Chapter I) is an inappropriate goal. Rather, he suggests that researchers and evaluators should stop trying to justify this federal funding program and focus instead on ways in which compensatory education practices can be improved. Thus, Cooley and others are arguing that we should not be asking whether programs like Chapter I make a difference, but rather how we can adjust or align the programs so that they do make a difference. As noted above, we have learned a great deal in the recent past about the characteristics and features of educational programs that work. It is now time to adopt certain of the features that have been shown to be effective in compensatory education environments and to test others that have shown promise in non-compensatory settings. Thus, we turn now to a discussion of design features which must be considered in the delivery of

compensatory educational programs.

SETTING FOR INSTRUCTION

The first design feature we consider is the setting or location for instruction. Compensatory education can be delivered within the regular classroom, outside the regular classroom, or partially inside and partially outside the regular classroom. Instruction delivered outside the regular classroom has been given the name pullout instruction, since recipients are removed or "pulled out" of their regular classroom and sent to another location. Usually, the location is within the same school building, and the instruction delivered there usually is provided by an instructional specialist working with smaller groups of children than there were in the classroom from which the Chapter I student was removed. Instruction delivered within the regular classroom is called inclass or mainstreamed instruction. One approach to mainstreaming is for the regular teacher to instruct Chapter I students while Chapter I personnel, either a specialist teacher or an aide, provide supplementary assistance. Another form of mainstreaming occurs when a Chapter I staff person instructs Chapter I students while the regular teacher

works with the other students in the class. A third, albeit less popular strategy, is known as the replacement model, which provides in one classroom all the services a Title I student receives in a given subject. Finally, there are what are known as add-on programs which result from an extension of the school year (i.e., summer programs), the school week or the school day.

What do we know about the popularity of these approaches? Carter (1984) reported that compensatory instruction was typically provided in a pullout setting, as did Vanecko, Ames, and Archambault (1980), Archambault and St. Pierre (1980), Kennedy (1978), and Stonehill and Anderson (1982). NIE (1977) reported that almost 75% of the students in Title I reading programs received their instruction in a pullout setting. About 41% of the compensatory language arts instruction was also delivered in a pullout setting, as was about 45% of the mathematics instruction. NIE also reported that about 24% of compensatory education students receive their regular instruction in groups comprised of only CE students. Glass and Smith (1977) claim that when the NIE pullout figures are corrected to eliminate pupils in 100% Title I eligible classrooms who do not need to be pulled out, the

pullout rates in all other classrooms rise to 84% for reading, 54% for mathematics, and 50% for language arts. Glass and Smith go on to argue that "when one considers further that pupils might be "pulled out" for one of these subjects and not the other, it is plausible to say that in classes not 100% Title I eligible "the practice of 'pull out' for compensatory teaching is nearly universal" (p.2).

More recently, Advanced Technology, Inc. (1983) reported that over 90% of the districts in their sample employed a pullout design exclusively or in combination with an inclass model; only 30% of the districts used an inclass design. Advanced Technology also reported, however, that pullout programs were being used less frequently than they had been in the past and that the use of inclass designs was on the rise, particularly in larger school districts (i.e., over 10,000 students). They found much less frequent use of the replacement model. Although this paper will not address the use of add-on programs, partly because the data on them is particularly skimpy, it is interesting to note that summer programs were available in over half (51%) of the 5,010 schools surveyed in the Sustaining Effects Study. As reported by Heyns (1986), larger schools are more likely to

conduct summer programs than smaller schools. Interested readers are referred to the Heyns paper for more details on summer programs. It should be noted, however, that she concludes that such programs are not unequivocally successful.

Impact of Setting on Achievement

Given the predominant use of the pullout and inclass modes of instruction, this discussion of the impact of setting on achievement will focus on these two models. In general, the literature leads to the following conclusions: (1) in some cases compensatory education in a pullout setting results in higher student achievement than inclass instruction, in other instances inclass instruction has a more positive effect, and in still others, and these appear to be the majority of cases, the setting for instruction has no significant effect on student achievement; (2) the findings on the cost effectiveness of the approaches are also mixed; and (3) the type of instruction delivered within a particular setting is more important than the setting itself.

One of the first large scale investigations of the effect of instructional setting for compensatory education students was the third year evaluation of

the Emergency School Aid Act (Coulson et al., 1977). In one part of this evaluation residual gain scores on individual student reading and math tests were regressed on 19 variables which were thought to be related to achievement. One of the 19 variables was the proportion of reading or math instruction received in a pullout environment. The 8,319 students participating in this research were drawn from two types of compensatory education programs and from grades 3, 4, and 5. Since both reading and mathematics scores were analyzed, a total of 12 regression equations were derived, one for each program, grade level and content area. For eight of these analyses, no significant differences were found for the pullout variable. In the other four instances, three of which were concerned with reading, pullout was found to have a significant negative effect on achievement. Leinhardt and Palley (1982) interpret these findings to mean that "a change from receiving no pullout instruction in reading to receiving one half reading instruction in pullout would be associated with a small reduction in the dependent measure" (p. 569). They also argue that no matter what the dependent measure was, the effect was very small.

Perhaps even more important than the small effect size is the strong possibility that pullout was associated with less gain because (a) less able students tended to get pullout instruction and (b) the pretest did not control adequately for such differences. Further, and as suggested by Coulson et al. (1977), schools were more likely to use pullout instruction when they had students with more severe problems.

The Instructional Dimensions Study (IDS) conducted for NIE by the Learning Research and Development Center (LRDC) at the University of Pittsburgh also was concerned with the effect of pullout. This research focused on reading and mathematics instruction in first and third grade and used classrooms as the basic unit of analysis. Here the variable "setting" was scaled as a composite of the time students were in supplemental instruction outside the regular classroom and the number of children in a classroom that were pulled out for instruction. NIE (1977) reported that the results uncovered in this research differed by content area and grade level. First grade students receiving instruction in the regular classroom (i.e., the inclass setting) made significantly larger gains in reading and mathematics than those in pullout

settings. However, at the third grade level setting had no significant effect on reading achievement and pullout was associated with larger achievement gains in mathematics. NIE interpreted these results to mean that neither setting is consistently associated with greater instructional effectiveness.

Frechtling and Hammond (1978) reported on a reanalysis of the NIE findings. These researchers, who were part of the team monitoring the IDS study for NIE, noted that rechecking and cleaning the data changed the previously described picture slightly. As before, inclass instruction was found to have a positive effect on achievement at grade 1 for both reading and mathematics. At grade 3, however, inclass instruction was found to be more effective for reading (there was no difference previously) while no significant difference was found for instructional setting when mathematics was considered (pullout was more effective previously) {see Table 1}. Frechtling and Hammond (1978) argued that despite these changes in the previously reported findings they would "be uncomfortable concluding from the IDS data that districts should use only the inclass approach" (p.9). They offered several reasons for their position: (a) the IDS findings were based on data

Table 1

Various Perspectives on the Pullout Findings of the Instructional Dimensions Study[1,2]

(1) NIE's Perspective

A

Grade	Reading	Math
1	I^+	I^+
3	ND	P^+

Based on NIE (1977)

B

Grade	Reading	Math
1	I^+	I^+
3	I^+	ND

Based on Frechtling & Hammond (1978)

(2) LRDC's Perspective

C

Grade	Reading	Math
1	ND	ND
3	I^+	ND

Based on Cooley & Leinhardt (1980)

D

Grade	Reading	Math
1	P^+	P^+
3	ND	P^+

Based on Leinhardt & Palley (1982)

[1]The Instructional Dimensions Study was commissioned and monitored by NIE and conducted by the Learning Research and Development Center at the University of Pittsburgh.

[2]Key: I^+ = inclass students outperform pullout students; P^+ = pullout students outperform in class students; ND = no difference.

from districts that were "special" and not nationally representative; (b) they had a gut feeling that mainstreaming (i.e., inclass instruction) may pose problems for many teachers and simply not be possible in some instances; and (c) they were unwilling to generalize from a single study. On this last point they argued that "the IDS, despite all its virtues, is only one study, and convergent evidence from other work is not yet strong enough" (p.9).

Cooley and Leinhardt (1980) and Leinhardt and Palley (1982) also report on the IDS study and, to make matters somewhat more confusing, their findings are different from those reported by NIE. Cooley and Leinhardt (1980) reported that setting was related to student achievement gain only in third grade reading and that more time in pullout was associated with less gain. As can be seen in Table 1, this is quite different from NIE's interpretation of the findings.

Leinhardt and Palley (1982) discuss a re-analysis of the IDS data in which they at least partially addressed a criticism by Glass and Smith (1977) that the pretest measure used was too difficult and thus did very little to adjust for initial group differences . In these analyses they included only children who had received pullout instruction and they

treated setting as a student level variable rather than the classroom level variable employed in their previous analysis. They also changed the nature of the setting variable, redefining it as the number of minutes a child was in pullout. This reanalysis led them to conclude the following:

> Reanalyzing the data with only those children that received pullout changed the picture dramatically. In the analysis with the non-pullout students, there was a negative relationship between the time spent in pullout and posttest in all cells. When the non-pullout students were excluded from the analysis, the results completely changed. There was a positive significant relationship between pullout and posttest in all cases except third grade reading. (p. 568)

By way of explaining differences across the two analyses, Leinhardt and Palley (1982) offer the following:

> The most plausible interpretation of these results is that the nonpullout students had more academic knowledge than the pullout students, that is, the poorer students were chosen for pullout. Hence, pullout was associated with less achievement

> because it had less able students and not because it was an inferior practice. Pretest did not adequately adjust for all the initial differences. In this case, the underadjustment would be particularly faulty for the first grade since (as Glass & Smith, 1977, pointed out) the pretest was not a good measure of initial abilities. Examining the results with the nonpullout students removed, we see the following: Pullout is still associated with less gain in third grade reading but not in the other three cases. Overall, the more time a student spent in the segregated pullout class, the greater the gain. It seems likely that effective practices were being implemented in these settings. It cannot be determined whether students who were pulled out suffered some initial losses due to having been removed from the regular class. We can say that once pulled out, they increasingly gained from more pullout instruction. (p. 568)

Prior to these reanalyses, Glass and Smith (1977) looked at the results of both the IDS study, as reported by NIE, and the ESAA evaluation and found that the ESAA study provided a better data base and

more believable evidence for assessing the effects of pullout. Their review of the ESAA findings led them to conclude that there was a consistent negative relationship between the percentage of time pupils spend in the pullout setting and their mathematics and reading achievement. They also concluded that "pupils pulled out of their regular classrooms would have to receive remarkably effective compensatory programs to offset the potential risks incurred. In our opinion, the 'pulled out' pupil is placed in moderate jeopardy of being dysfunctionally labelled, of missing opportunities for peer tutoring and role modeling, and of being segregated from pupils of different ethnic groups" (p. 41). Finally, they concluded that the "pulled out" procedure per se had no clear academic or social benefits and may, in fact, be detrimental to pupils' progress and adjustment to school.

From their paper, it appears that Glass and Smith were willing to base their conclusions about pullout on the results of a single study. Obviously, other studies influenced their decision, as suggested by their review of the impact of variables such as ability grouping, labeling, peer tutoring, racial desegregation, and mainstreaming the handicapped.

But, it appears that the ESAA findings were the sine qua non of their argument. With these findings in hand they could argue that pullout is not only ineffective in improving achievement but that it also produces harmful labeling of students, it serves as a means for resegregating students, and so forth. But is pullout truly ineffective in improving achievement? The reanalysis of the IDS data argue that it is not, and, based on these findings, it appears that Glass and Smith (1977) may have been too negatively disposed to the impact of pullout, at least on achievement. But, at the same time, the reanalyzed IDS data do not suggest that pullout is the best method for improving the achievement of compensatory education students. Rather, they seem to suggest that pullout is one method that should be considered. Cooley (1981) captures the essence of these mixed reactions: "No doubt there are schools in which the pullout practice is being done well, but there are certainly schools in which this may be doing more harm than good" (p. 300).

More recently, a few studies have looked at the impact of pullout on achievement, but, with one exception, these studies are much narrower in scope and in significance than those reported above.

Consequently, they do little more than add fuel to the setting debate fire. Nonetheless, they are worthy of consideration, and we turn now to a brief discussion of them.

The Sustaining Effects Study did not directly address the question of the impact of pullout on achievement. However, Carter (1984) did conclude that pullout was the most typical mode of instruction in Title I programs and that Title I did make a difference in achievement, if only for moderately academically disadvantaged students. It could be argued that Carter thus indirectly concluded that pullout has some effect on achievement. However, and as is increasingly the case in the literature, Carter was more interested in describing the effect that particular classroom practices had on achievement than in discussing the effect of setting. Those variables which his research found to be important were opportunity to learn, lower student/staff ratio, allocation of teachers' time, classroom management practices, source of student off-task behavior, supervision of students, and coordination of instruction. He also found years of teaching experience and the instructional leadership of the principal to be important variables.

Yap (1983) also addressed the question of pullout in his secondary analysis of data acquired over a three year period (1978 to 1981) throughout the State of Hawaii. In this research Yap compared pullout programs with both inclass and combined inclass/pullout programs. These comparisons were performed separately for reading and mathematics, and for reading only by year, using analysis of covariance procedures in which an NCE gain score was the dependent variable, type of program was the independent variable, and pretest NCE and per pupil cost were covariates. Yap offers no explanation for his selection of per pupil cost as a covariate. Moreover, although he reports sizeable differences in per pupil cost (e.g., almost $350 more per pupil in 1978-79 for pullout versus inclass instruction), he does not describe how this covariate effects the dependent measure, if at all, and it is not possible to determine from the results whether it did. Yap also does not explain why he chose to use the more unreliable gain score as the dependent variable, particularly given the use of the NCE pretest score as a covariate.

Yap reported that there were some 300 "projects" involved in his research. That is, 287 schools were

involved in the reading analysis and 26 were involved in the mathematics analysis. For reading, 75 schools were involved in 1978-79, 109 schools in 1979-80, and 103 schools in 1980-81, and in all years over 60% of them were pullout schools. Yap bases his major conclusions on school level comparisons. This means that NCE's had to be aggregated across grades. It also means that only one type of setting must have been employed at a given school. Yap offers no guidance as to whether this did occur.

Clearly, Yap's study had some weaknesses, and, consequently, his results are somewhat suspect. Nonetheless, it is interesting to note that Yap found no negative achievement effects resulting from pullout. In fact, his school-level analyses reveal a trend indicating the superiority of pullout programs over both inclass and combined pullout/inclass programs. Given small sample sizes and missing data problems, Yap reported that he was unable to ascertain through a formal analysis of variance or covariance whether grade level interacted with setting. However, an inspection of unadjusted and adjusted means indicated to Yap (no means were presented) that, although such effects were not substantial, there may be grade level differences. Taken together these

results lead Yap to conclude that in terms of achievement gain, students who received Chapter I instruction in the pullout setting are likely to perform as well as, if not better, than their counterparts in other settings. He also concluded that, despite some drawbacks, the pullout setting was a viable option for providing services to Chapter I youngsters.

In a 1985 paper, Yap also argued for the cost effectiveness of the pullout approach. Using the data described above, with all of the inherent weaknesses ascribed to them, he found that pullout produced a unit of NCE gain for every $81 spent. Corresponding figures for inclass and combination programs were $84 and $86, respectively. These findings disagree with Fitzgerald and Hunt (1985). In this study of about one third of the school districts in Minnesota, pullout and inclass settings were found to be equally effective in producing student achievement gains. However, the pullout setting was found to be more costly and to show a trend toward lower cost-effectiveness than inclass instruction. Fitzgerald and Hunt argue that these increased costs might be due to the use of licensed teachers in pullout settings and teacher aides for supplemental

inclass instruction. Given the scope of this study (141 randomly selected districts) and the quality of the methods used throughout, it is possible that the Fitzgerald and Hunt findings on cost effective are closer to reality than those reported by Yap.

Nearine and Pecheone's (1984) study of the Intensive Reading Instruction Team (IRIT) Program, which was implemented in Hartford, Connecticut and validated by the State of Connecticut and the JDRP, also reported positive effects for pullout. This study included 698 students in grades 3 through 6 whose pre-entry reading performance on the Metropolitan Achievement Test fell below the 23rd percentile. The comparison group contained 124 third grade through sixth grade students who also scored below the 23rd percentile but who were not selected for the program (i.e., a matched sample). Since students entered this 10-week intensive program at different times during the school year, comparisons could also be made across three "cycles." Thus, post-test performance of students completing earlier cycles could be compared with pre-test performance of students entering a later cycle. Nearine and Pecheone reported gains of at least 6 NCEs across the various reading and language arts analyses. They also

reported that students who completed an IRIT cycle performed higher than students entering the next cycle. Finally, they found that the gains in reading were sustained throughout the school year (i.e., beyond the time of IRIT instruction). However, they did not find signficant post-test differences between IRIT students and those in the comparison group, who outperformed the IRIT students on the pre-test, although not significantly.

Knight (1979a, 1979b, and 1979c) also reported positive effects of pullout in studies of several schools in New York City. In each of these studies the actual post-test scores of pullout participants were compared against the scores students would be expected or predicted to receive without intervention. In all of the comparisons performed across the three studies, actual post-test performance exceeded predicted levels. Further, all but 6 of the more than 25 comparisons across various grade levels and content areas were statistically significant. However, the weaknesses of the research design and alpha problems associated with multiple comparisons render these findings somewhat suspect.

Finally, Madhere (1981) reported that pullout was more effective than inclass instruction in improving

student achievement. Madhere also argued that pullout appeared to be most effective in an environment where the Title I pupils consititute a majority of the students in the school. He contended that in such a context, the impact of labeling a student in need of assistance, a negative consequence according to the critics of pullout, may be meaningless.

Doss and Holley (1982), on the other hand, reported that pullout programs were not effective. They argued that compensatory programs conducted during the regular day inevitably supplanted regular instruction. They also argued that one of the reasons for the ineffectiveness of pullout programs was the regular classroom teacher's decreased sense of responsibility for the special program student. Based on Glass and Smith's (1979) meta analysis of class size and achievement, they argued that reducing class size to 15 and delivering instruction within the regular classroom would not only redress some of the problems occurring within pullout programs, but would also improve student performance. Since this approach was permissible under the "schoolwide projects" provision of Title I in force at that time, such a program was implemented on a pilot basis in 1980-81 in two Austin, Texas elementary schools. Analyses were

conducted on reading and mathematics achievement for students in grades 2 through 6 in those schools and in certain Title I schools using pullout programs. From these analyses, Doss and Holley concluded that in most schoolwide projects students outscored Title I regular students of the same pretest level and that the impact of the schoolwide projects was generally uniform across all levels of the pretest. On average, the advantage of the schoolwide projects across grades were 2.1 months in reading, 2.5 months in language arts, and 2.2 months in mathematics. However, these gains were purchased at such great expense ("lowering the PTR (Pupil Teacher Ratio) from 25-1 to 15-to-1 would increase personnel costs by 67% in the schools and grades where applied" (Doss and Holley, 1982, p.10)) that the project had to be discontinued after about two years of operation. Levine and Stark (1982) report that the schoolwide projects implemented in the Los Angeles Unified School District suffered a similar fate partly because they were very expensive and partly because there were questions about how well they worked.

What, then, can we say about the effect of pullout on achievement? Based on the research that has been conducted to date, we cannot conclude with confidence

either that pullout is more effective than inclass instruction or that the opposite is true. We can conclude, however, that the achievement findings provide little support for the overwhelming, and in some locations almost exclusive, use of the pullout model. Rather, the findings suggest that both pullout and inclass instruction can be effective, given appropriate circumstances, and, as reported by Carter (1984), Leinhardt and Palley (1982), Tobias (1982), and others, it is likely that these circumstances, particularly the type and quality of the instruction delivered within an instructional setting, are more important than the setting itself. Despite this general sense, however, it must be reported that pullout programs have been opposed for reasons other than their impact on achievement.

Other Effects of Instructional Setting

In addition to achievement findings, the literature also includes an increasing amount of other information on the effects of inclass versus pullout programs. Included among the long list of issues which have been addressed are the following: (a) the stigma associated with receiving Chapter I instruction, particularly in a pullout setting; (b) the levels of communication, cooperation and

coordination possible in the two settings; (c) the levels of teacher autonomy and control afforded by the settings; (d) the amount and type of content lost when compensatory instruction is received, particularly in pullout programs; (e) class size shifts and resultant gains; (f) grouping patterns, individualization and segregation in inclass and pullout settings; (g) the relationship among peers and the opportunity for peer tutoring; and (h) the types of instructors and their roles in different settings. The first four of these issues will be addressed in this section. The remaining four issues will be addressed in sections which follow. In either case, much of what we know about these issues is derived from relatively small scale observational studies, studies of relatively modest scope reporting on the results of interviews with program students, teachers, and administrators, and somewhat narrow surveys of the perception of these and other groups. Moreover, as was the case with the review of the impact of setting on achievement, the data which are available are generally equivocal.

Stigma. Leinhardt and Palley (1982) reported that despite contrary research findings, special educators strongly believe that special placement makes the target child feel rejected rather than rescued.

Likewise, many educators believe that compensatory education students feel different if they are labeled as Chapter I eligible and are pulled out of their regular classrooms for special instruction (e.g., Leinhardt and Palley (1982); Shuy (1978); Glass and Smith (1977)). Some argue, on the other hand, that stigmatization occurs in mainstream programs as well (Noddings, 1978). On this score, at least one author has suggested that stigmatization may be greater in the mainstream setting since in this environment compensatory education students are both singled out and remediated in front of their classmates (Hayes, 1983). From a slightly different perspective, it has also been argued that pullout is an effective instructional strategy in an environment where Chapter I students constitute a majority, and that in such settings labeling may be meaningless (Madhere, 1981).

Even though the research evidence is inconclusive, it appears reasonable to argue that stigmatization occurs in both inclass and pullout settings: in the inclass setting because students receive extra instruction in front of their peers, and in the pullout setting because students are easily identifiable. It also appears that it is not the setting itself that aggravates the feeling of being

different, but the sensitivity or insensitivity of teachers operating within the setting (Noddings, 1978). Kennedy's (1978) advice on this issue seems particularly compelling:

> Compensatory education students are less likely to be subjected to labeling and its negative effects in environments where teachers actively encourage children's respect for and appreciation of a variety of human differences; in environments where similarities between learning tasks and materials are emphasized; and in environments where Title I and non-Title I children frequently move, in an organized way, to other parts of the building or classroom to receive special instruction. (p.35)

Communication, Cooperation, and Coordination. One of the most frequently cited advantages of the inclass setting is the increased communication, cooperation, and coordination among regular classroom and compensatory education instructors that it affords. Advocates of inclass instruction argue that because specialist teachers or aides hired with compensatory education funds provide their instruction in the same classroom as the regular teacher, there will be increased interaction and cooperation in planning for

instruction (Bean and Eichelberger, 1985; Kennedy, 1978; Hayes, 1983; Neumann, 1985), better integration of remedial and regular instructional strategies (Allington, 1986; Bean and Eichelberger, 1985; Kaestle and Smith, 1982), better opportunities for individualization and small group instruction (Bean and Eichelberger, 1985), and a better flow of information on student progress (Harnischfeger, 1980). There will also be fewer disruptions caused by students moving in and out of the regular classroom (Hayes, 1983), and thus more continuity in instruction, as well as an increased opportunity for regular classroom teachers to observe other, perhaps more effective, approaches for working with compensatory education students (Kennedy, 1978).

Allington (1986) has offered some particularly powerful comments on one of the consequences of pullout, namely the fragmentation of instruction. He suggests that pullout remedial reading instruction is typically independent from the reading instruction offered in the regular classroom and that it generally employs different materials and teaching methods. He argues further that neither the classroom teacher nor the compensatory education specialist is likely to know what type of instruction the other is offering.

Consequently, he claims that few remedial students actually receive instruction that supplements their core curriculum; rather they are taught by classroom and remedial teachers who express different beliefs about student needs, offer different objectives as targets for instruction, and use reading material that represent different models of the reading process. As a result, he argues, students who are in need of remediation are offered instruction that seems likely to increase their confusion rather than reduce it. Allington also acknowledges, however, that pullout programs can offer instruction that is congruent with that provided in the regular classroom (i.e., same or similar material, same reading strategy used and same reading skill taught), and, therefore, that the lack of congruence may be related to factors other than the instructional setting itself.

Despite this indictment of pullout programs, it appears that inclass programs also present problems, some of which are the same as those prevalent in pullout settings. Bean and Eichelberger (1985), for example, noted differences in classroom teachers' and specialists' teaching styles and instructional strategies when a large city school district changed from a pullout program to one in which specialists

worked in the regular classroom. Moreover, although teachers in this new program reported increased interaction and cooperation, they also reported problems with scheduling and inadequate joint planning time; reading specialists, on the other hand, reported that there was a lack of teacher interaction, that they weren't sure what their role was in the classroom, that they had insufficient materials, and that there was inadequate management. Bean and Eichelberger (1985) concluded that these problems were due to two teachers sharing the same room and that, despite the inclass arrangement, team teaching was difficult. Vasquez and Nuttall (1983) reached similar conclusions. These researchers also argued that it is the classroom teacher, not the specialist, who is forced to schedule activities that will accommodate another class in the same room and that the coordination between the regular and compensatory program is still accomplished outside of the regular classroom, before school, during recess, during lunch, or after school. They argued further that even though the materials used in the regular and remedial programs are more likely to be similar in inclass programs, pullout programs use a greater variety of instructional materials. Specialist teachers working

in the regular classrooms apparently find it difficult to transport books and other paraphernalia from classroom to classroom (Bean and Eichelberger agree), find it harder to use audiovisual materials because they would be distracting to the regular classroom teacher, and, for that same reason, avoid materials that require recitations or other "out-loud" work. Despite his enthusiasm for the inclass approach, Neumann (1985) concluded that where this type of setting is used specialists must be highly organized, must be flexible enough to work in several environments, and must have the diplomacy and personality needed to work jointly with colleagues.

Taken together, these results suggest that even though inclass programs offer the opportunity for increased communication, cooperation, and instructional continuity, they do not insure that they will occur. Perhaps, as suggested by Kennedy (1978), the key to good coordination between compensatory and regular instruction is not the setting for instruction but rather the quality of the joint planning that occurs, the amount of decision-making power given to teachers and specialists by their respective administrators, and the flexibility in scheduling. Whether these characteristics occur more frequently in

inclass settings remains to be seen.

Teacher Autonomy and Control. Critics of the pullout model have argued that when students receive remedial instruction outside the regular classroom neither the regular teacher nor the compensatory education specialist is willing to accept full responsibility for the students' progress (Allington, in press; Doss and Holley, 1982). If they so desire, classroom teachers can claim that less than acceptable progress is attributable to shortcomings of either the student or the compensatory instruction; compensatory education specialists, on the other hand, can attribute such a result to either the student or inadequate regular classroom instruction. To avoid such academic fingerpointing, some have argued that the adoption of the inclass model would allow the teachers sharing the same space to share this responsibility or the administrators coordinating the program to decide on whose shoulders the responsibility should rest. Despite the wisdom of this argument, however, it appears that the adoption of an inclass arrangement will again not insure the expected outcome. Bean and Eichelberger (1985), for example, reported that both teachers and specialists had problems with two instructors in one room and that

the problem of leadership or control was a real issue. Vasquez and Nuttall (1983) reached a similar conclusion, as did Shuy (1978), who found that some of the teachers in the Instructional Dimensions Study favored pullout because it gave them greater autonomy. On the other hand, Shuy also reported that some teachers favored the inclass approach because it gave them more control. Unfortunately, the literature does not allow one to resolve this apparent conflict.

Content Lost. It was argued earlier in this paper that students receiving compensatory instruction in a particular content area (e.g., reading) do so at the expense of regular instruction in that same content. Harnischfeger (1980) concluded that pullout pupils "often receive less reading time than the rest of the class because the pulling out - going some place and coming back - also entails loss of instructional time for students who need more" (p. 4). Regardless of whether this is true, and there is some evidence to suggest that the total time allocated to basic skills instruction is greater for compensatory education students than for their non-compensatory education peers (Archambault and St. Pierre, 1980), the implication is that some approach other than or in addition to pullout would be more effective.

Harnischfeger's solution would be to lengthen the school day for compensatory education students (i.e, an add-on approach). Others argue that providing inclass instruction would solve the problem because then the time lost by traveling from place to place would be retained. Allington (1986), for example, suggests that as much as 15 minutes per day, or forty hours per academic year, might be saved in this manner. However, neither Allington nor anyone else that this reviewer has been able to uncover provides direct evidence on what would actually be done with the newly found time were compensatory instruction to be provided inclass. Would the interruptions prevalent in inclass programs (Hayes, 1983) prove to be less troublesome than the disruption associated with pullout programs? Would the presence of more than one teacher in a single classroom result in greater confusion than occurs in pullout programs, particularly when such programs are adopted on a wholesale basis? Additional research is needed before we can determine with any confidence whether inclass compensatory education programs result in more or less time allocated to instruction. Additional research is also needed to determine whether such differences in allocated time are accompanied by differences in the

amount of time which students in various settings actually engage or actively participate in instruction.

Class Size, Individualization, and Grouping

Compensatory instruction is delivered in smaller classes and groups than regular classroom instruction. It is also more likely than regular instruction to incorporate components of individualized instructional programs. Pullout instruction, which comprises the large portion of compensatory instruction, is delivered in groups that are more homogeneous in ability than those in the classrooms from which compensatory students are removed. Do the reductions in class and group size and the increase in individualization of instruction have a positive affect on students? Is homogeneous ability grouping preferred to heterogeneous ability grouping? We turn now to a discussion of these issues.

Glass and Smith (1979) reported on the results of a meta-analysis of 77 studies involving 725 comparisons of larger and smaller classes. They found that 60% of the achievement comparisons favored smaller classes and that the effect was particularly

strong when the classes contained fifteen or fewer pupils. They also found that this relationship did not change significantly for students of different ability levels. On the other hand, a review of 24 studies conducted by the Educational Research Service (1978) concluded that research provided no clearcut guidelines for an optimum class size. The ERS review also concluded, however, that pupils with lower ability tend to benefit from smaller classes more than pupils of average ability and that smaller classes can positively affect the scholastic achievement of economically or socially disadvantaged pupils. Cooper (1986), in his comprehensive review of the literature, summarizes the effect of class size on achievement as follows:

> Both the (Glass and Smith) meta-analysis and the ERS review concluded that smaller class sizes benefit the population of students served by Chapter I programs. Glass and Smith did so by finding the benefit of small classes held for students regardless of their intelligence level. ERS found the benefit for low ability or economically disadvantaged children only. Further, the typical size of compensatory education instructional groups falls within the

> range about which there appears to be some agreement. It will be recalled that from one-third to all of Title I students in the districts sampled by the (Title I) Demonstration Study (Archambault & St. Pierre, 1979) received compensatory instruction in groups of one to five students. All the reviewers agreed that this is the end of the class size curve where effects are most dramatic and dependable. (pg. 9)

Hedges and Stock (1985) cited two problems with Glass and Smith's estimation procedures. First, they argued that they overestimated effect sizes when study Ns were small. Second, they argued that the variance of a student's effect size estimate was not based on its sample size, as it should have been. Hedges and Stock (1983) thus reanalyzed the same class size data taking these problems into account. As might be expected, they found a smaller effect size than Glass and Smith. However, they also found that smaller classes still led to higher expected achievement and that the shape of the curve expressing the relationship between class size and achievement

found by Glass and Smith remained unchanged.

Slavin (1984) has criticized both the Glass and Smith study and the Hedges and Stock reanalysis. He argues that the Glass and Smith results were heavily influenced by studies of tutoring (i.e., one-to-one instruction, or a class size of one). This claim is based on his review of 14 studies listed by Glass and Smith in which students were randomly assigned to large or small groups. Fourteen of the 30 contrasts in these studies involved comparisons of classes with a median size of 23 to those with class sizes of one to five, and 10 of these involved classes with one student. The median effect size for these comparisons was .545. Nine involved comparisons of classes with a median size of 35 to classes ranging in size from 14 to 17. The median effect size for these studies was .08. Finally, for comparisons involving classes with more than 20 students (i.e., median larger classes of 37 and median smaller classes of 26), the median effect size was .04. Thus, Slavin concludes that the effects of tutoring are positive but that "there is nothing in the Glass et al. meta-analysis to contradict an assertion that class size makes no meaningful difference in achievement unless the smaller class has no more than three students" (p.

11).

Slavin's criticism of Hedges and Stock (1983) focuses on their study selection procedures. That is, although he agrees with Hedges and Stock's conclusion about overestimated effect sizes, he believes that Hedges and Stock's reanalysis produced the same general conclusion as Glass and Smith because they chose to include the same studies as the earlier researchers.

The differences in the conclusions reached by these analysts seem to hinge on the assumptions which guide their work. Glass et al. (1981) argue that all studies that meet certain selection criteria, including poorly designed studies, should be included in a meta-analysis. Slavin (1984, 1986) argues that only good quality studies should be included in a review of the literature, and that those of higher quality should receive more weight. Thus, his analysis of a smaller number of rigorous comparisons (i.e., 30 comparisons in 14 studies employing random assignment) leads Slavin to argue that Glass and Smith are not able to support their class size conclusions. Despite the plausability of Slavin's arguments, there is general agreement in the literature that smaller classes, particularly those with fewer than 15

students, result in higher student achievement. Moreover, and as argued by Cooper (1986), because the strongest effects are found with class sizes in the 1-5 range, these small class sizes must be considered a benefit for compensatory education students.

The Glass and Smith study also looked at the effect of class size on variables other than achievement. They concluded that teacher morale is higher, that their attitude toward pupils is more positive, and that their satisfaction with their own performance is greater in smaller classes. They also found positive effects of smaller classes on "student attitude", which included attitudes toward the teacher, school, and themselves (i.e., self-concept), as well as positive effects on student motivation. Cooper (1986) summarizes these other results of the Glass and Smith study:

> This category of effects [student attitudes] demonstrated a difference of 29 percentile ranks between students in classes of size 40 and size 10. Another group of effects involving measures of student participation showed a difference between classes of 10 and 40 students equal to 23 percentile ranks. Finally, a group of variables measuring different types of student misbehaviors

also favored smaller classes, but the number of comparisons was too small to estimate a percentile rank difference. (p. 11)

Individualization

Since the average class size for compensatory instruction in reading, language arts, and mathematics are all below 14 while the average size of the regular classroom is 27 (Archambault & St. Pierre, 1979; Carter, 1984; NIE, 1977), it would appear from the above that smaller class sizes would be preferred. Moreover, if students in these smaller classes are also receiving more individualized instruction, as Glass and Smith (1980) say they do, then it would be logical to conclude that class sizes for compensatory education students, whether the instruction is compensatory or not, should be small (i.e., at least less than 15), and that individualized instruction should be emphasized. Before we leap to this conclusion, however, we should review what other researchers have to say.

The Class Size and Instruction Program (CSIP) research of Cahen et al. (1983), which included four second grade classes selected from an innercity school in California and a rural school in Virginia,

investigated what happens in classrooms when class size is reduced. These researchers concluded that smaller classes made discipline easier and that teachers in them spend more time teaching and less time policing. They also found that students paid closer attention in smaller classes and that group discussions involved more of the students in the group. Finally, the researchers concluded that the effect of participation may be most pronounced for low achievers because "in a small group, where control is perceived to be easier, the teacher may feel she or he can take time to draw all students into the lesson" rather than "rely on volunteers or high ability students in order to keep things moving along." (cited in Cooper, 1986, p. 13).

The Class Size and Instruction Program as well as the Sustaining Effects Study (Carter, 1984) reported on behavior management issues associated with the sizes of classes and groups within them. The Sustaining Effects Study found that the smaller instructional groups and lower student-to-instructor ratios associated with Title I programs were associated with more student on-task behavior, less teacher time in behavioral management, a more harmonious class environment, and a higher quality of

cognitive monitoring, task monitoring and organization of activities. The CSIP researchers reached the same general conclusions.

CSIP also looked at the relationship between class size and individualization. Although the CSIP researchers recognized that smaller classes provided the opportunity for greater individualization of instruction, they found that smaller classes did not lead to dramatic increases in the amount of individualization. Instead, they found that group instruction continued and the curriculum remained the same. Harnischfeger (1980) concluded that, even when class sizes are reduced drastically, many teachers do not significantly change their teaching strategies. She suggests that the only consequence of smaller class sizes is smaller subgroups but not fewer groups, and, on the average, no increased teacher/pupil interaction. Harnischfeger goes on to argue that instead of a teacher working with small groups one at a time while the rest of the class is on their own, he or she should consider whole class instruction. If they do, Harnischfeger argues, their students will benefit by receiving more direct instruction.

Despite the evidence provided by the CSIP researchers, one might argue that the findings are

suspect because of the small sample of schools and classrooms (i.e., 2 schools and 4 classrooms involved in the research) and that individualization does occur more frequently in smaller classes. Suppose this were true? What do we know about the impact of individualization on student outcome measures?

Cooley and Leinhart (1980) reported they did not find any clear evidence for the superiority of individualized instruction over other methods of compensatory education. Their review of the literature also suggested that this finding is consistent with other attempts to establish the effectiveness of individualization. Brophy (1986) has argued that the problem is not with the conceptual model of individualization, which calls for beginning where students are and moving them along at their own pace. Rather, the problem is that:

> Individualized instruction in the typical classroom shifts a great deal of responsibility for planning and managing learning from the teacher to the students themselves and shifts responsibility for carrying the content to the students from the teacher to the materials. This is workable and may even have certain advantages when the teacher is continually available to

> provide close supervision and immediate help when needed, but it does not work well when students must function on their own for extended periods of time, trying to learn through interacting with the curriculum materials without much guidance or help from the teacher. (pp. 13-14)

Brophy does state, however, that the individualized training package approach can be used effectively in special classrooms with small student/teacher ratios. Basing his conclusion on the results of a study by Crawford (1983), Brophy indicates that under such conditions "teachers can move the students through curriculum at a faster pace, can provide more tutorial and individualized instruction, and can assign more difficult work because they are able to monitor everyone's progress and provide immediate help when needed" (p. 14). Nonetheless, Brophy's review underscores the role of active instruction from the teacher in producing achievement gains:

> Such instruction can be provided in tutorial form in special classes with small student/teacher ratios, and it can be provided to small groups in classrooms where the teacher has developed appropriate assignments and installed effective seatwork management procedures so that students

> are engaged in worthwhile academic activities in between their small group lessons from the teacher. The individualized approach is not feasible in the typical classroom, however, and even the small group approach is likely to strike most teachers as more trouble than it is worth unless they have an aide available to supervise seatwork while they teach small group lessons (or unless they believe that the class is so heterogeneous that they are forced to provide differentiated instruction to small homogeneous groups). Consequently, most teachers will opt for the traditional whole-class instruction/recitation/seatwork method as their primary approach to instruction. (pg. 15)

What does Brophy mean by active teaching? Active teaching connotes frequent lessons (whole class or small group, depending on grade level and subject matter) in which the teacher presents information and develops concepts through lecture and demonstration, elaborates this information in the feedback given following responses to recitation or discussion questions, prepares the students for follow-up assignments by giving instructions and going through practice examples, monitors progress on assignments,

and follows up with appropriate feedback and reteaching when necessary. There is a great deal of teacher talk, but most of it is academic rather than procedural or managerial, and much of it involves asking questions and giving feedback rather than extended lecturing.

Grouping

The issue of grouping for instruction has been alluded to several times in the preceding discussion, but to this point we have said little or nothing about how and why groups are formed and the effects of different grouping patterns. In general, ability grouping refers to the practice by which students in more heterogeneous classes or groups are arranged into more homogeneous classes or groups for all of a portion of their instruction. Most American schools follow this practice, and most teachers overwhelmingly support it, probably because they believe that homogeneous groups are easier to teach than heterogeneous groups. Dividing students according to ability allows teachers to increase the pace and level of instruction for high achievers and provide more repetition, review and remediation for low achievers. Ability grouping is also supposed to provide a spur to

high achievers by making them work harder to achieve, and to place success within the grasp of low achievers, who are protected from having to compete with their more able agemates (Atkinson & O'Connor, 1963). Critics argue, however, that ability grouping confers unnecessary distinction on those in the fast track while placing a stigma on those in slow groups (Kulik & Kulik, 1982). They also contend that being placed in a low achieving class or track has damaging psychological effects on students (Oakes, 1985; Rosenbaum, 1980; Schaefer & Olexa, 1971).

Slavin (1986) in his "best evidence synthesis" of the literature on ability grouping in the elementary school carefully distinguishes within-class from between-class ability grouping. Within class ability grouping refers to the assignment of students to more homogeneous groups within a heterogeneous class arrangement. The most common form of within class grouping is the use of reading groups, where teachers assign students to one of a small number of groups on the basis of reading level. These groups work on different materials at rates unique to their needs and abilities. Other forms of within-class ability grouping may be found in mathematics, where the teacher presents a lesson to the class as a whole, and

afterwards, while the students are working problems, the teacher provides enrichment or extension to a high-achieving group, remediation or re-exploration to low achievers, and something in between to average achievers. According to Slavin, group-paced mastery learning (Bloom, 1976) may be seen as one form of flexible within-class ability grouping. Individualized or continuous-progress instruction may be seen as extreme forms of ability grouping.

Slavin (1986) also contends that between-class groups takes many forms. In fact, he describes the following seven (7) types:

(1) Ability Grouped Class Assignment - Students are assigned on the basis of ability or achievement to one self-contained class (usually at the elementary level) or to one class which moves together from teacher to teacher, as in block scheduling in junior high schools.

(2) Curriculum Tracking - A special form of ability grouped class assignment unique to the secondary level in which students are assigned by ability or achievement to tracks, such as college preparatory, general, or vocational. Students may take all or only

some of their courses within these tracks.

(3) Regrouping for Reading or Mathematics (Ability Grouping for Selected Subjects) - Students are assigned to heterogeneous homeroom classes for part or most of the day, but are "regrouped" according to achievement level for one or more subjects.

(4) Joplin Plan - A special form of regrouping for reading in which students are assigned to heterogeneous groups for most of the day but are regrouped for reading across grade levels (Floyd, 1954). For example, a reading class at the fifth grade, first semester reading level might include high-achieving fourth graders, average achieving fifth graders, and low-achieving sixth graders. Reading group assignments are frequently reviewed, so that students may be reassigned to a different reading class if performance warrants it.

(5) Nongraded Plans - In its broadest application, grade-level designations are entirely removed, and students are placed in flexible groups according to their performance level, not their age (Goodlad & Anderson, 1963). Students move from group to

group and complete the curriculum at their own pace so that some students might complete the primary grades, for example, in two years while others would require four years.

(6) Special Classes for High Achievers - Plans in which gifted, talented or otherwise superior students are assigned to a special class for all or part of their school day, while other students remain in relatively heterogeneous classes.

(7) Special Class for Low Achievers - the assignment of students with learning problems to special or remedial classes for part or all of their school day.

Dreeban (1984) suggests that factors other than a student's ability or past achievement also affect their assignment to groups. Included among these additional factors are the distribution of students' aptitudes or achievement in the school or classroom, organizational constraints within the school or classroom, and student management and discipline. With regard to assignment within classes, teachers also consider the number, size and diversity of students within groups when assigning other students to the group. Generally, teachers form at

least three groups, and groups formed on the basis of ability may vary widely on other dimensions. Nonetheless, instruction is geared to the level of ability attributed to the group.

With regard to within-class grouping, a recent review of the literature by Wilkinson (1986) has concluded that students in low ability reading groups fare poorly:

> Students in low groups receive less time to read in comparison with students in high groups; oral reading errors made by lows are more frequently corrected than those made by highs (i.e., they are interrupted more frequently); lows are given more instruction in decoding and teacher prompts are more focused on phonics for lows than for highs; lows are frequently less attentive than highs (Allington, 1980; Eder, 1981, 1982, 1983, 1984; Rist, 1970; Stern & Shavelson, 1981). Eder (1981) argues that these practices result in lower reading-test performance and lower oral-reading grades for the low ability groups. The poor achievement of the lows is thus the result of low expectations by the teachers and/or the poor quality of teaching that they receive. At first glance, these studies suggest that the

> differential treatment received by the low ability students is the culprit, the cause of their poor reading achievement. (pp. 12-13)

Wilkinson continues:

> The implications for low-achieving students seem clear: Homogeneous ability grouping is detrimental to learning for students assigned to low groups (Eash, 1961, Persell, 1977). These studies reveal mechanisms through which the negative effect of assignment to a low ability group occurs. There are differences in the instructional processes and learning environments of different ability groups. Teachers and students interact in different ways across ability groups. We are compelled to agree with Hallinan's (1984, p. 233) conclusion that: `At present, the great majority of studies demonstrate that assignment to a low track or ability group places students at a disadvantage in terms of learning opportunities.' (p. 13)

Wilkinson also concluded that low ability students seem to benefit achievement-wise by being placed in small, heterogeneous ability groups (Swing & Peterson, 1982) and that these groups can provide environments that are positive for learning (Cohen, 1984; Webb,

1982, 1985; Webb & Kenderski, 1984; Peterson, Wilkinson, Spinelli, & Swing, 1984).

Slavin's (1986) review of the literature on within-class ability grouping reached a different conclusion. Basing this portion of his review exclusively on research in mathematics (because no reading studies met stated criteria for inclusion - see Slavin, 1986, pp. 15-16 for a discussion of these criteria), Slavin found that every one of the eight studies he located favored homogeneous within-class grouping, though not always significantly. Five of the studies also broke down achievement effects by student ability levels, but there was no clear pattern of results favoring ability grouping for one or another subgroup. On this score Slavin concluded the following:

> Every subgroup gained more in classes using within-class ability grouping than in control (ungrouped) treatments. However, it is interesting to note that the median effect size for low achievers (ES = +.65) was higher than that for average (ES = +.27) or high achievers (ES = .41). (p. 84)

Slavin acknowledges that this evidence is surprising in light of several studies of ability grouping in

reading suggesting that students in low reading groups experience a lower quality of instruction than do those in higher groups.

How can we explain the discrepant findings from the Wilkinson and Slavin reviews? One possible explanation is that Wilkinson draws heavily on studies in reading whereas Slavin focuses only on mathematics. More important than this, however, is the fact that much of the research cited by Wilkinson deals with observations that students in low reading or math groups learn less than students in higher groups, not less than other low-achieving students in contrasting arrangements. Thus, according to Slavin, it is hardly surprising that high and low achievers differ and that their teachers' behavior differs accordingly. Slavin argues further that:

> Comparisons of achievement gains in high and low reading groups are found to show an advantage of being in the high group because high achievers learn more rapidly than low achievers, and unless measures used to control for initial ability are perfectly reliable and perfectly predictive of later reading achievement, assignment to the high reading group will appear to lead to higher achievement (see Reichardt, 1979)" (Slavin, 1986,

p. 58).

What then can we conclude about the effect of within-class ability grouping? At present, the best evidence in this area comes from Slavin's review, and, based on this, it seems safe to conclude that at least in mathematics and on achievement grounds, within-class ability grouping should be supported for all students, including low achievers. However, and despite the results presented by Wilkinson, it does not appear reasonable to conclude at this time that either homogeneous or heterogeneous within-class ability grouping is preferred for delivering reading instruction to low ability students. As suggested by Slavin, such a conclusion can be reached only when experimental comparisons of grouped and ungrouped reading classes have been performed.

Turning now to between-class ability grouping, Glass and Smith (1977) called the research literature on ability grouping "a confusing welter of contradictory, tendentious, antiquated and dubious studies" (p. 32). Nonetheless, drawing from the reviews of Miller and Otto (1930), Billet (1932), Cornell (1936), Ekstrom (1959), Eash (1961), and Goldberg, Passow & Justman (1966), Glass and Smith (1977) concluded that ability grouping is not an

important variable in itself, and that it is worth considering only as it relates to or potentiates other changes in instructional activities. They also concluded that "the accumulated research literature of several decades shows such as unstable pattern of what takes place between teachers and pupils in homogeneous as opposed to heterogeneous classrooms that no generalizations are possible (or useful) about the effects of ability grouping" (pp. 34-35).

Kulik and Kulik (1982) conducted a more recent meta-analysis of the effects of ability grouping on secondary school students. They found that in 36 of the 51 studies looking at achievement, the performance of students in grouped classes was better than the performance of students in ungrouped classes, in 14 other studies the performance was better in ungrouped classes and in one study there was no difference. A total of 10 of the comparisons were statistically significant, eight of which favored grouping and two of which favored what Kulik and Kulik called "conventional instruction." The effect size was small, however, averaging about one-tenth of a standard deviation across studies. In the four programs designed especially for academically deficient students the effects were even smaller. In

fact, they were near zero, indicating that students learned as much in mixed ability classrooms as they did in homogeneous classrooms. Kulik and Kulik did find, however, that homogeneous grouping was effective for high achievers.

Slavin (1984) criticized the Kulik and Kulik (1982) review on the grounds that many of the studies they included used matching or other non-random procedures to attempt to equate groups on initial ability, and that this procedure would result in selection bias. Moreover, Slavin argued that in studies of high achievers, this bias should favor homogeneous assignment for gifted students, as was found by Kulik and Kulik. Slavin also argued that since the mean effect size for studies using random assignment is zero, Kulik and Kulik's claim of positive effects of ability grouping rests entirely on nonrandomized studies.

In his own review of the ability grouping literature, Slavin (1986) argues that research clearly does not support the assignment of students to self-contained classes according to ability, nor does it support the conclusion that ability grouping is beneficial for high ability students and detrimental for low ability students, as suggested by Eash (1961),

Esposito (1973), and Begle (1975). The literature does suggest, however, that cross-grade tutoring and comprehensive non-graded plans can be effective. Particularly noteworthy in this regard are the fourteen studies reviewed by Slavin using Joplin, or Joplin-like, Plans. For eleven of these studies the Joplin plan classes achieved more than the control classes; for the remaining three studies no differences were found. In addition, the median effect size for these studies was found to be about .45. The results for the comprehensive non-graded plans are less consistent than for the Joplin-like plans, but the median effect size is still moderately positive (ES = .29). Based on this, Slavin concludes that the evidence supports the use of these plans. He also suggests that while the results of studies of regrouping for reading and mathematics are inconclusive, it is still quite possible that grouping for one or two subjects is instructionally effective.

Related to the literature on ability grouping is an increasingly large body of research on the relative effects of mainstreaming versus special class placement. Due in large part to the impact of the 1975 Education of All Handicapped Children Act (PL 94-142), which mandated that handicapped children had

to be educated in the least restrictive environment, most of the recent research has focussed on this special population, particularly the mildly handicapped, who prior to 1975 had most often been served in self-contained settings completely separate from the public schools. Because the results of these studies are not directly generalizable to the types of students served by compensatory education programs, these mainstreaming findings must be interpreted cautiously. However, it should also be recognized that the findings are not without merit and may be applicable to those now served by Chapter I and other state and local programs for the academically disadvantaged. In fact, as suggested by Leinhart, Bickel and Palley (1982), it may well be that some treatments are effective for both the mildly handicapped and the academically disadvantaged.

Leinhart and Palley's (1982) informative review of the literature on mainstreaming and restrictive educational settings concluded that while there is a trend toward negative effects of isolated settings for low ability children, it is neither strong nor highly consistent. They also concluded that children on the borderline of the cutoff for special placement suffer more when they are isolated than do children who are

closer to the median of the group of children that receive special services. Finally, they concluded that the setting itself (i.e., mainstreaming versus isolated setting) is not a powerful direct inference on student achievement, but that it indirectly influences other relevant variables, most of which can occur in most settings.

Like Leinhart and Palley (1982), Madden and Slavin (1983) also concluded from their review of the literature that full-time special education placements resulted in few consistent benefits for the mildly academically handicapped (i.e., learning disabled or educable mentally retarded). However, Madden and Slavin went further by concluding that regular class placement (i.e., mainstreaming) was preferred to special class placement on achievement grounds. The best evidence for this conclusion, Madden and Slavin argue, is the research of Calhoun and Elliott (1977) in which 50 black educable mentally retarded (EMR) students and 50 black emotionally disturbed students were randomly assigned to regular and special classes and followed over a three year period. These researchers found that placement in regular classes had a more positive effect on the achievement of both groups than did special class placement.

Interestingly, Madden and Slavin (1983) also cite a study by Leinhardt (1980) as support for regular class placement. In this study

> Low achieving students in the regular class made significantly greater gains in reading achievement than did students in the special transition rooms, despite the fact that the transition rooms, which used the same curriculum, had much lower student teacher ratios and were staffed by some of the schools `best' teachers. (Madden & Slavin, 1983, p. 524)

Madden and Slavin admit that their conclusions are based heavily on these two studies. However, they argue that their conclusion is warranted since the two studies are particularly well designed and their effects are quite large.

Despite his general conclusion that regular class placement is preferred for the mildly academically handicapped, Slavin suggests that the "effects of regular class placement depend substantially on the program in the regular class" (1986, p. 530). Specifically, he argues that individualized instruction is required for the positive effects to occur, and that this applies to social-emotional outcomes as well as achievement. He also suggests,

however, that part-time resource programs can have a positive impact on the achievement of the mildly academically handicapped if they are of good quality, again hinting that this implies that individualized instruction occurs. In this regard, Slavin seems to be at odds with Brophy (1980) and others who are calling for less individualized instruction and more direct or active instruction.

Despite his enthusiasm for regular class placement, Slavin recognizes that the placement of mildly academically handicapped students in regular classes is fraught with problems. Included among these is the fact that regular class teachers feel poorly prepared to accommodate the needs of the mildly academically handicapped and that LD and EMR students in regular clases are less accepted and more rejected than their non-handicapped peers. Slavin describes a variety of programs for overcoming these problems, and proposes that these programs and strategies be used to either restructure the regular classroom or to construct resource programs to supplement instruction in the regular class. Thus, in the language of ability grouping, Slavin appears to be calling for heterogeneous groupings of students, including the mildly academically handicapped, within-class. He

also suggests, however, that some form of regrouping for certain types of instruction, probably employing pullout since he uses the resource room terminology, may be required.

But even if pullout programs can be justified on academic grounds, removing children from regular classrooms for compensatory instruction creates the possibility of resegregation. On this score, Glass and Smith (1977) have argued that pullout programs do not alter the racial mixture of schools but they may and probably do change the racial composition of the classes within then. Eyler (1982) is less equivocal. Citing the work of Hinkley, Beal, and Breglio (1978) and Brookover, Brady, and Warfield (1981), she argues that minority students receive larger amounts of compensatory instruction than non-minorities and that they are typically pulled out of less segregated classrooms and sent to more segregated special rooms to receive their instruction. Citing the work of Kimbrough and Hill (1981), Eyler argues that the potential for resegregation through compensatory services is exacerbated in schools that operate more than one categorical program and have substantial numbers of students who are eligible for more than one type of service. She also argues that

the negative effects of resegregation are not offset by pullout's effects on achievement. To avoid problems created by pullout, Eyler recommends that compensatory services be delivered inclass by specialists who act as consultants rather than simply as subject matter specialists. In this redefined role they would serve as resources to the regular classroom teacher rather than instructors of specific groups of students. They would assist teachers is assessing specific learning problems and preparing individual learning plans, train classroom aides and parent volunteers, and help the principal plan the schoolwide instruction program.

On the question of race relations, Slavin (1979, 1980) and Slavin and Madden (1979) reported that working together on school projects with a student of another race and competing together on interracial sports teams both had strong positive effects on racial attitudes and behaviors. However, teacher workshops, minority history, multiethnic texts, bisocial student advisory committees, and similar school programs made no difference in racial attitudes or interracial friendships. Given these results, it would appear that if pullout programs do lead to resegregation, as Eyler (1982) and others have argued

that they do, the resultant reduction in contact between minority and non-minority students will reduce the possibility of improved relations among the races. To the extent that this occurs, it must be viewed as a serious negative consequence of homogeneous grouping which uses the pullout model.

Type of Instructor

As noted earlier in this review, instructional specialists provide the large majority of compensatory instruction. Generally, these specialists have a higher level of educational attainment than regular classroom teachers. In general, they also have less experience, and, because of this, they are also likely to receive less pay. Despite the widespread use of these better trained but less experienced teachers in compensatory education programs, the literature tells us surprisingly little about the effect which they have on students. Glass and Smith (1977) report on two studies which claim to show that remedial reading specialists have a positive effect on reading performance (Flynn, Hass, Al-Salam, 1976; Kiesling, 1971). However, neither this review nor the one reported here was able to uncover any additional information on this issue. Perhaps because of this

lack of information, Lignon and Doss (1982) have suggested that compensatory funds be used to purchase additional classroom teachers rather than reading specialists, thereby reducing the size of the regular classes. Allington (in press), on the other hand, believes there is ample support for the continued use of reading specialists. He does feel, however, that reading specialists will be under increasing pressure to justify their positions, particularly to their classroom teacher peers.

Concerning the nature of their services, Eyler (1982) has suggested that the role of compensatory instructional specialists should be redefined to require them to work in inclass settings and serve as resources to the classroom teacher rather than as instructors of specific groups of students. Interestingly, Bean and Eichelberger (1985) have reported that reading specialists perceived that they placed less emphasis on both diagnosis of student weaknesses and teaching specific skills and more emphasis on working with content teachers when their school district changed from a pullout to an inclass approach. Classroom teachers concurred with specialists and also reported that specialists spent more time working with them. Despite the similarity

between what Eyler (1982) proposed and what actually occurred in the Bean and Eichelberger (1985) study, the latter authors reported that teachers and specialists both agreed that reading specialists should focus on working with students with reading problems. Neither group saw the reading specialist as a resource to the classroom teacher, although both groups valued having interaction about specific children.

Shifting our attention to instructional aides, Schuetz (1980) reports that the evidence about their effectiveness is mixed. Qualitative studies have found that teachers using instructional aides spend more time on instructional activities, deliver more of their instruction in individual and small groups, feel they accomplish more, and feel that instructional aides are helpful. Quantitative studies of their impact on student achievement, however, are less conclusive. In fact, Schuetz reports positive, negative, and no effects attributable to instructional aides in her comprehensive review of the literature. Despite the lack of convincing evidence for their effectiveness, Schuetz provides useful information on the conditions which must be satisfied if aides are to be helpful: (1) teachers who use instructional aides

must be genuinely receptive to them; (2) instructional aides must have the necessary literacy and computational skills to perform some useful instructional tasks in the classroom; (3) classroom management plans must be developed to take advantage of the presence of instructional aides; and (4) instructional aides must receive intensive training in the tasks they will be asked to perform.

Viewed as a whole, it is clear that the literature concerning the impact on student performance of specialist teachers and instructional aides is inconclusive. This may not be surprising, since by this time in the review most readers may have grown accustomed to equivocal results. What may be surprising, however, is the lack of evidence about the effectiveness of instructional specialists and instructional aides regardless of the type of instructional program or activity in which they are involved. Perhaps Allington (in press) is right in arguing that reading specialists will be called on increasingly to justify their positions. Whether he is or not, research clearly needs to be performed to determine exactly what impact reading specialists have on student performance and whether the continued use of compensatory educations funds to purchase their

services is warranted.

Summary and Conclusions

Since the passage of the Elementary and Secondary Education Act of 1965 billions of dollars have been spent each year to attempt to provide improved instruction for the poor and educationally disadvantaged. Despite the magnitude of this investment and the commitment of the many thousand people who have labored to insure that these monies are well spent, the impact of compensatory education, although measurable and significant (Carter, 1984; Forbes, 1985, Stickney & Plunkett, 1982; Wang, 1980) has nonetheless been viewed as disappointing by some (e.g., Hallinan, 1986) and less than hoped for by others (e.g., Cooley, 1981). Cooley (1981), however, has also argued that the search for overall effects of compensatory education is an inappropriate goal. Rather, he suggests that researchers and evaluators should stop trying to justify this federal funding program and focus instead on ways in which compensatory education can be improved. With this goal in mind, a review of the literature was undertaken to determine what is known about the effectiveness of certain design features of programs

for the educationally disadvantaged. Specifically, the review was concerned with the effects of particular instructional settings (i.e., whether the instruction was delivered inclass or outside the regular classroom in a pullout setting), class size, instructional grouping practices, individualized instruction, and instructional specialists and aides.

Although recent research indicates that the inclass approach is being used more frequently (Advanced Technology, 1983), the literature clearly shows that most compensatory instruction is delivered in a pullout setting. It also shows that compensatory instruction is predominately reading and mathematics instruction, and that it is provided primarily by instructional specialists, and to a lesser extent instructional aides who work with smaller groups of students and provide more individualized instruction than regular classroom teachers.

The research describing the impact of pullout programs on the achievement of compensatory education students is less clear. In fact, based on what we know at this time, one would be unable to provide convincing evidence for the effectiveness of either the pullout or inclass approach. Given these mixed results, however, one could and should argue that the

widespread use of pullout programs is not justified. Moreover, since there is some evidence to suggest that pullout may result in resegregation, there is further reason to question its wholesale adoption. But that does not mean that pullout programs should be abandoned. Rather, it means that the programs implemented should fit local conditions.

Critics of pullout have suggested that it unncessarily labels and stigmatizes students, that it reduces communication, cooperation, and coordination between regular and compensatory instructors, that is reduces teacher responsibility for pupil progress, and that valuable content is lost when pullout instruction is received. Despite the importance of these criticisms, counterarguments also appear to have some value. Labeling and stigmatization can occur in inclass settings as well as in pullout settings. Moreover, it may be that it is not the setting itself which aggravates a compensatory education student's feelings of being different, but the sensitivity or insensitivity of teachers operating within the setting (Noddings, 1978). It also appears that even though an inclass program makes it easier and more likely for communication, cooperation, and coordination to occur, it does not guarantee that they will. There is also

some evidence to suggest that certain types of teachers are just not able to work together in a single classroom and that joint planning may take place in much the same way in inclass programs as in pullout programs, before school, during recess, during lunch, and after school. Because some teachers may not be able to accommodate another teacher in the same classroom, the issue of teacher autonomy and responsibility for student progress is also not automatically resolved by inclass programs. Perhaps, as suggested by Kennedy (1978), the key to good coordination between compensatory and regular instruction is not the setting for instruction but rather the quality of the joint planning that occurs, the amount of decision-making power given to teachers and specialists by their administrators, and the flexibility in scheduling allowed by the environment.

With regard to the content lost issue, it is clear that some instruction is sacrificed when pullout compensatory instruction is received. It is also clear that valuable instructional time is lost in the process of moving to and from the pullout environment. it is not clear, however, what types of instruction, if any, would be lost if compensatory instruction were delivered inclass. It is also not

clear whether the interruptions resulting from students being pullout out for instruction are less harmful than the confusion resulting from two teachers working in the same classroom. More research needs to be done on these issues.

Those who favor pullout programs herald the virtues of specialist teachers and aides, smaller class sizes, increased opportunities for individualization, and more effective grouping practices. What do we know about their effectiveness? Surprisingly, we know very little about the impact of instructional specialists and aides on student performance. In fact, this review has been able to uncover only two studies that claim to show that reading specialists have a positive effect on achievement (Flynn, Hass, Al-Salam, 1976; Kiesling, 1971). Allington (1986) believes there is ample support for the continued use of reading specialists, but his arguments are not particularly convincing. Moreover, he cautions reading specialists that in the future they will be under increasing pressure to justify their positions, particularly to their classroom teacher peers. Regarding classroom aides, Schuetz (1980) reports that quantitative studies lead to no clear conclusions about their

effectiveness. She does suggest, however, that some teachers can use them effectively, but that in order for this to occur, teachers must be genuinely receptive to them, the aides must have the necessary literacy and computational skills to perform useful instructional tasks, they must receive intensive training in the tasks they will perform, and effective classroom management plans must be developed to accommodate them.

Despite the paucity of hard, convincing evidence about the effectiveness of instructional specialists and aides, they continue to be used as the primary agents of compensatory instruction. In the early days of Title I their use was clearly justified given the legal requirements surrounding the allocation of these federal monies. In fact, it has been argued (Archambault, 1986) that these requirements were precisely the reason that pullout programs, with the accompanying specialists and aides, became so popular. At present, however, there are fewer constraints. And yet, even the inclass models which are being implemented use instructional specialists or aides as the service providers. We need more research on their effectiveness; we need more creative approaches to their use.

Turning now to the class size issue, with the exception of the work of Slavin (1984) there seems to be general agreement that class sizes of less than fifteen optimize learning outcomes. Moreover, and as argued by Cooper (1986), the effects are most dramatic in classes containing from 1 to 5 students, the size of the group in which the majority of compensatory education students receive their instruction (Archambault and St. Pierre, 1979). Thus, it would appear that smaller class sizes are to be favored. But what of the composition of the groups within these classes? Is heterogeneous or homogeneous grouping preferred?

Slavin's (1986) comprehensive review of the effects of instructional grouping strategies offers the best organization of the literature on this subject as well as some of the best insights into what that literature says. Slavin distinguishes within class and between class ability grouping, and he considers the effects of each separately. He concludes that homogeneous within-class ability grouping leads to improved student performance in mathematics and that this effect holds for students of all ability levels. This conflicts with the conclusions of Eash (1961), Pursell (1977), Wilkinson

(1986), and Hallinan (1986) that homogeneous ability grouping is detrimental to students assigned to low groups. One explanation for this difference is that Slavin focused on mathematics while the other reviewers, particularly Wilkinson (1986), focused more on reading. More important than this, however, is the fact that much of the research cited by Wilkinson deals with observations that students in low reading or math groups learn less than students in higher groups, not less than other low achieving students in contrasting arrangements. Like Reichardt (1979), Slavin argues that when groups are not equivalent pre-experimentally (i.e., when there is no random assignment to groups) such comparisons will tend to favor higher achieving groups.

Taken together, then, the research suggests that for mathematics within-class ability grouping should be supported for all students, including low achievers. However, it does not appear reasonable to conclude at this time that either homogeneous or heterogeneous within-class ability grouping is preferred for delivering reading instruction to low ability students. As suggested by Slavin, such a conclusion can be reached only when experimental comparisons of grouped and ungrouped reading classes

have been performed.

Turning now to between-class ability grouping, a wealth of advice is available in the literature. Some reviewers have concluded that no generalizations are possible about the effects of ability grouping (Glass & Smith, 1977), others have concluded that while the performance of students in grouped classes is better than the performance of students in ungrouped classes there appears to be no systematic effect for low ability students (Kulik & Kulik, 1982), and still others have argued that ability grouping is detrimental for low ability students (Begle, 1975; Earl, 1971; Esposito, 1973). Slavin (1986), who looked at the effects of seven types of between-class ability grouping strategies, concluded that research does not support the assignment of students to self-contained classes according to ability, nor does it suggest that ability grouping is harmful for low ability students. It does suggest, however, that cross-grade tutoring and comprehensive non-graded plans can be effective. It also suggests to Slavin that while the results of regrouping for reading and mathematics are inconclusive, it is still possible that grouping for one or two subjects is instructionally effective.

The last of the practices which the present review addressed was individualization. This would appear to be a particular benefit for compensatory education students since the reduced class size in which they receive their compensatory instruction should result in more individual attention. However, Cahen et al. (1983) and Harnischfeger (1980) have both concluded that smaller class size does not necessarily lead to increased individualization. Cooley and Leinhardt (1980) have also concluded that there is not any clear evidence for the superiority of individualized instruction over other instructional strategies. Brophy (1986) argues that the problem is not with the conceptual model of individualization, which calls for beginning where students are and moving them along at their own pace. Rather, the problem is that as implemented individualized instruction shifts the responsibility for teaching to the student. This does not work when used over extended periods of time. It can be used effectively, however, in classrooms with small student teacher ratios, where the teacher is able to provide close supervision and where other types of instructional methods are used in conjunction with individualized instruction, particularly what Brophy calls "active instruction."

What do these results suggest for practice? What advice can we offer to those charged with the implementation of compensatory education programs? Under the best of conditions the translation of research into practice is a risky business. When the evidence is not clear-cut the perils are even greater. Researchers need to recognize this. They also must recognize that practitioners want and need help and that the practice of schooling cannot wait for unambiguous research results. With this in mind, and with no little amount of trepidation, the following advice is given:

1. Setting - Based on what we know at this time, it is safe to conclude that setting is not directly responsible for student outcomes and that the issue of effective practices within settings, and not setting itself, should be the focus of our attention. This implies that educators can more confidently select a setting based on local conditions, perhaps using pullout approaches in those areas where resegregation would not occur, where teachers are committed to the use of these practices, where coordination and cooperation is possible, and so forth. But, where these and certain other conditions are not met, and perhaps

even within the same school district, inclass programs should be used.

2. Specialists and Aides - It is clear that classroom teachers need help in meeting the needs of the academically disadvantaged. It is less clear whether that help should be provided by instructional specialists, aides or some other type of teacher. Since there is so little evidence telling us what combinations of instructors and conditions are likely to maximize student performance, the best advice is that school districts should again feel confident in using whatever combination appears to meet local needs. In certain situations it may be best to have specialists work with students in a pullout setting; in others it may be more effective to introduce a specialist or aide into the regular classroom. If the latter option is chosen, it may be better in some situations for the specialist to work with compensatory education students; in others the skills of the teacher and specialist may dictate that the specialist should work with other students or groups while the classroom teacher concentrates on the needs of the compensatory education student. In determining

which option is best, school administrators should honestly assess the competencies of their staff and attempt to match them to the needs of students. They should also consider the needs of teachers and their preferences for and their ability to adapt to the various options. Finally, and no less importantly, they should determine which of the options is likely to deliver the highest quality instruction to the student. Key issues in this determination should be which program will insure the greatest amount of engaged time in specific subject areas, the best communication, cooperation, and coordination among the regular and compensatory program, the greatest integration of instruction across these two domains, and the fewest negative consequences such as the stigma attached to receiving compensatory instruction.

3. Class Size and Grouping - Smaller class sizes are preferred, particularly when they include fifteen or fewer students. Educationally disadvantaged students benefit from such arrangements, as do other students in the class. Educationally disadvantaged students also benefit from classes which are heterogeneous in ability, and they do so

at no expense to the higher ability students. This does not deny the advantages of within-class ability grouping for certain subjects, particularly reading and mathematics, nor the possibility that cross grade grouping for certain content would also be beneficial. In fact, both of these practices are to be encouraged.

4. Individualization - Individualized instruction can be harmful when used as the sole method of instruction. That is not to say that teachers should not determine the individual needs of students and help them to progress at their own pace. It does mean, however, that teachers should not transfer the responsibility for this progress to students laboring on their own for extended periods of time or worksheets and other published materials. Said in a different way, teachers must take a much more active role in instruction than the practice of individualization has come to imply. Individualization can be used effectively, however, in classrooms with small student teacher ratios where the teacher is able to provide close supervision.

From a number of quarters and for a number of reasons effective educational practices are now being

emphasized in the literature. Edmonds (1979) and Brookover et al. (1978) talk about effective schools, that is, schools in which there is a commitment to particular instructional models and programs, there is agreement on goals and objectives, principals assume a leadership role in the determination of what instruction is to be offered and how it is delivered, administrators and teachers have high expectations for students and communicate these expectations to them, there is frequent monitoring and feedback of student performance, and there is effective teacher training. Leinhardt and Palley (1982), looking more closely at classrooms, call for smaller classes, mastery learning (i.e., clear goal statements and regular monitoring of student progress), more time spent in cognitive activities, reasonably rapid pacing of instruction, using a formal management system which encourages task orientation and successful completion of tasks, and increased teacher instructional time. Brophy (1986), echoing some of Leinhardt and Palley's comments, would use this increased time for instruction for what he calls active teaching. This connotes frequent presentations by the teacher in which she develops concepts through lecture and demonstration, follows those up with recitations by students or dialog with

them, prepares the students for assignments by giving instructions and going through practice examples, monitors progress on assignments, and follows up with appropriate feedback and reteaching when necessary. There is a great deal of teacher talk, but most of it is academic rather than procedural or managerial, and most of it involves asking questions and giving feedback rather than extended lecturing. These and other researchers have taught us a great deal about what makes schools and the people within them effective. Nonetheless, we have much more to learn.

REFERENCES

Advanced Technology, Inc. (1983). Local operation of Title I: A resource book. Washington, DC: U.S. Department of Education.

Allington, R. L. (1986). Policy contraints and the effective delivery of remedial instruction. In J. Hoffman (Ed.), Effective teaching of reading: Research and practice. Newark, DE: International Reading Association.

Archambault, F. X. (1986, June). Instructional setting: Key issue and bogus concern. Paper presented at the Office of Educational Research and Improvement sponsored and Research and Evaluation Associates, Inc. conducted Conference on the Effects of Alternative Designs in Compensatory Education, Washington, DC.

Archambault, F. X. & St. Pierre, R. G. (1980) The effect of federal policy on services delivered through ESEA Title I. Education Evaluation and Policy Analysis, 2, 33-46.

Atkinson, J. W. & O'Connor, P. (1963). Effects of ability grouping in schools related to individual differences in achievement-related motivation. Final Report, Cooperative Research Project No. 0E-2-10-024. Washington, DC: U. S. Department of H.E.W.

Bean, R. T. & Eichelberger, R. T. (1985). Changing the role of reading specialists: From pullout- to in-class programs. Reading Teacher, 38(7), 648-653.

Begle, E. G. (1975). Ability grouping for mathematics instruction: A review of the empirical literature. Stanford Mathematics Education Study Group, Stanford University. (ERIC Document Reproduction Service No. ED 116 938).

Breglio, V. J., Hinckley, R. H., & Beal, R. S. (1978). Students economic and educational status and selection for compensatory education. (Technical Report. No. 2). Santa Monica, CA: System Development Corporation.

Billet, R. O. (1932). The administration and supervision of homogeneous grouping. Columbus: The Ohio State University Press.

Bloom, B. S. (1976). Human characteristics and school learning. New York: McGraw-Hill.

Brookover, W. B., Brady, N. M., & Warfield, M. (1981). Educational policies and equitable education. A report of studies of two desegregated school systems. In Procedures and pilot research to develop an agenda for desegretation studies. East Lansing, MI: Michigan State University, College of Urban Development, Center for Urban Affairs.

Brookover, W. B., Schweitzer, J. H., Schneider, J. M., Beady, C. H., Flood, P. K., & Wisenbaker, J. M. (1978). Elementary school social climate and student achievement. American Educational Research Journal, 15, 301-308.

Brophy, J. (1986, June). Research linking teacher behavior to student achievement: Potential implications for instruction of Chapter I students. Paper presented at the Office of Educational Research and Improvement sponsored and Research and Evaluation Associates, Inc. conducted Conference on the Effects of Alternative Designs in Compensatory Education, Washington, DC.

Cahen, L. S., Filby, N. N., McCutcheon, G., & Kyle, D. W. (1983). Class size and instruction. New York: Longman.

Calhoun, G. & Elliott, R. (1977). Self concept and academic achievement of educable retarded and emotionally disturbed pupils. Exceptional Children, 44, 379-380.

Carter, L. F. (1984). The sustaining effects study of compensatory and elementary education. Educational Researcher. 1984, 21, 4-13.

Cohen, E. (1984). Talking and working together: Status, interaction and learning. In P. Peterson,

L. C. Wilkinson, & M. Halliman (Eds.), The social context of instruction: Group organization and group processes. Orlando: Academic Press, 171-188.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). Equality of educational opportunity. Washington, DC: Government Printing Office.

Cooley, W. W. (1981). Effectiveness in compensatory education. Educational Leadership, 38, 298-301.

Cooley, W. W. & Leinhardt, G. (1980). The instructional dimensions study. Educational Evaluation and Policy Analysis, 2(1), 7-25.

Cooper, H. M. (1986, June). Chapter I programs reduce student-to-instructor ratios but do reduced ratios affect achievement? Paper presented at the Office of Educational Research and Improvement sponsored and Research and Evaluation Associates, Inc. conducted conference on the effects of Alternative Designs in Compensatory Education, Washington, DC.

Cornell, E. L. (1936). Effects of ability grouping determinable from published studies. In G. M. Whipple (Ed.), The ability grouping of pupils. NSSE Yearbook, 35, Part I. Bloomington, IL: Public School Publishing Co., 289-304.

Coulson, J. E., Ozenne, D. G., Hanes, S. D., Bradford,

C., Doherty, W. S., Duck, G. A., & Hemenway, J. A. (1977). The third year of the Emergency School Aid Act (ESAA) implementation. Santa Monica, CA: System Development Corporation. (ERIC Document Reproduction Service No. ED 154 952)

Crawford, J. (1983). A study of instructional processes in Title I classes: 1981-82. Journal of Research and Evaluation of the Oklahoma City Public Schools, 12(1).

Doss, D. & Holley, F. (1982, April). A cause for national pause: Title I schoolwide projects. Paper presented at the Annual meeting of the American Educational Research Association, New York, NY. (ERIC Document Reproduction Service No. Ed 214 996)

Dreeban, R. (1984). First-grade reading groups: Their formation and change. In Peterson, P., Wilkinson, L. C., & Hallinan, M. (Eds.). The social context of instruction: Group orgranization and group processes. Orlando: Academic Press, 69-84.

Eash, M. S. (1961). Grouping: What have we learned. Educational Leadership, 18, 429-434.

Edmonds, R. A. (1979). A discussion of the literature and issues related to effective schooling. St. Louis, MO: Central Midwestern Regional Educational Laboratory.

Educational Research Service (1978). Class size: A summary of research. Arlington, VA: Educational Research Service.

Ekstrom, R. B. (1961). Experimental studies of homogeneous grouping: A critical review. School Review, 69, 216-226.

Esposito, D. (1973). Homogeneous and heterogeneous ability grouping: Principal findings and implications for evaluating and designing more effective educational environments. Review of Educational Research, 43, 163-179.

Eyler, J. (1982). Resegregation: Segregation within desegregated schools. (ERIC Document Reproduction Service No. ED 216 075).

Fitzgerald, N. B. & Hunt, R. K. (1985, April). Cost-effectiveness predictors in Chapter I programs: The Minnesota example. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Floyd, C. (1954). Meeting children's reading needs in the middle grades: A preliminary report. Elementary School Journal, 55, 99-103.

Flynn, D. I., Hass, A. E., & Al-Salam, N. A. (1976). An evaluation of the cost effectiveness of alternative compensatory reading programs. Bethesda,

MD: RMC Research Corporation.

Forbes, R. H. (1985). Academic achievement of historically lower-achieving students during the seventies. Phi Delta Kappan, 66(8), 542-544.

Frechtling, J. A. & Hammond, P. A. (1978, March). Policy implications of the Instructional Dimensions Study. Paper presented at the annual meeting of the American Educational Research Association, Toronto, Ontario.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). Meta-analysis of social research. Beverly Hills, CA: Sage.

Glass, G. V. & Smith, M. L. (1977). 'Pull out' in compensatory education. Boulder, CO: University of Colorado, Laboratory of Educational Research.

Glass, G. V. & Smith, M. L. (1979). Meta-analysis of the research on class size and achievement. Educational Evaluation and Policy Analysis, 1, 2-16.

Goldberg, M. L., Passow, A. H., Justman, J. (1966). The effects of ability grouping. New York: Teachers College Press, Columbia University.

Goodlad, J. I. & Anderson, R. H. (1963). (Rev. ed.). The non-graded elementary school. New York: Harcourt, Brace, & World.

Hallinan, M. T. (1986, June). Chapter I and student

achievement: A conceptual model. Paper presented at the Office of Educational Research and Improvement sponsored and Research and Evaluation Associates, Inc. conducted Conference on the Effects of Alternative Designs in Compensatory Education, Washington, DC.

Harnischfeger, A. (1980). Curricular control and learning time: District policy, teacher strategy, and pupil choice. Educational Evaluation and Policy Analysis, 2(6), 19-30.

Hayes, M. V. (1983). A survey of teacher attitudes towards in-class versus pull-out compensatory reading programs. (ERIC Document Reproduction Service No. ED 228 619)

Hedges, L. U. & Stock, W. (1983). The effect of class size: An examination of rival hypotheses. American Educational Research Journal, 20, 63-85.

Heyns, B. (1986, June). Summer programs and compensatory education: The future of an idea. Paper presented at the Office of Educational Research and Improvement sponsored and Research and Evaluation Associates, Inc. conducted Conference on the Effects of Alternative Designs in Compensatory Education, Washington, DC.

Hinkley, R. H., Beal, R. S., & Breglio, V. J. (1978).

Student economic and educational status and receipt of educational services. (Technical Report No. 3). Santa Monica, CA: System Development Corporation.

Kaestle, C. F. & Smith, M. S. (1982). The historical context of the Federal role in education. Harvard Educational Review, 52, 383-408.

Kennedy, G. C. (1978). School setting and learning. In Perspectives on the Instructional Dimensions Study: A supplemental report from the National Institute of Education (pp. 33-39). (ERIC Document Reproduction Service No. ED 174 728)

Kiesling, H. J. (1971). Input and output in California compensatory education projects. Santa Monica, CA: Rand Corporation.

Kimbrough, J. & Hill, P. (1981). The aggregate effects of federal education programs. Santa Monica, CA: Rand Corporation.

Knight, M. E. (1979a). Evaluation of diagnostic/ prescriptive reading program, Community School District 31, New York City. (ERIC Document Reproduction Service No. ED 187 792)

Knight, M. E. (1979b). Evaluation of Title I program, Community School District 31, New York City. (ERIC Document Reproduction Service No. ED 187 791)

Knight, M. E. (1979c). Evaluation of Title I program,

Community School District 31, New York City. (ERIC Document Reproduction Service No. ED 186 544)

Kulik, C. C. & Kulik, J. A. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. American Educational Research Journal, 19(3), 415-428.

Leinhardt, G. (1980). Transition rooms: Promoting maturation or reducing education. Journal of Educational Psychology, 72, 55-61.

Leinhardt, G., Bickel, W., & Palley, A. (1982). Unlabeled but still entitled: Toward more effective remediation. Teachers College Record, 84(2), 391-422.

Leinhardt, G. & Palley, A. (1982). Restrictive educational settings: Exile or haven? Review of Educational Research, 52(4), 557-578.

Levin, H. M., Glass, G. V., & Meister, G. R. (1984). Cost-effectiveness of four educational interventions (Project Report No. 84-A11). Washington, DC: National Institute of Education.

Levine, D. U. & Stark, J. (1982, April). Instructional and organizational arrangements and processes for improving academic achievement at inner city elementary schools. A study of the Chicago Mastery Learning Reading Program and other school-wide

approaches for improving reading at selected schools in Chicago, Los Angeles, and New York. Paper presented at the annual meeting of the American Educational Research Association, New York, NY. (ERIC Document Reproduction Service No. ED 213 814)

Ligon, G. D. & Doss, D. A. (1982, April). Some lessons we have learned from 6,500 hours of classroom observations. Paper presented at the annual meeting of the American Educational Research Association, Montreal.

Madden, N. A. & Slavin, R. E. (1983). Mainstreaming students with mild handicaps: Academic and social outcomes. Review of Educational Research, 53(4), 519-569.

Madhere, S. (1981). Issues in program evaluation. (ERIC Document Reproduction Service No. ED 217 098)

McLaughlin, M. W. (1975). Evaluation and reform. The Elementary and Secondardy Education Act of 1965: Title I. Cambridge, MA: Bellanger Publishing.

Miller, W. S. & Otto, H. J. (1930). Analysis of experimental studies in homogeneous grouping. Journal of Educational Research, 21, 95-102.

National Assessment of Educational Progress (1981). Has Title I improved education of disadvantaged

students? Evidence from three national assessments of reading. Denver, Colorado.

National Institute of Education (1977). The effects of services on student development. Washington, DC: The National Institute of Education, U.S. Department of Health, Education, and Welfare.

National Institute of Education (1978). Perspectives on the Instructional Dimensions Study. Washington, DC: The National Institute of Education, U.S. Department of Health, Education, and Welfare.

Nearine, R. J. & Pecheone, R. L. (1984, April). Multiple assessments of a compensatory reading program over time. Paper presented at the annual meeting of the Amerian Educational Research Association, New Orleans, LA.

Neumann, T. (1985). In-class remedial instruction. An alternative to the pull-out programs. Wisconsin State Reading Association Journal, 29(3), 26-27.

Noddings, N. (1978). Overview of the Instructional Dimensions Study Conference. In Perspectives on the Instructional Dimensions Study. (pp. 5-14). (ERIC Document Reproduction Service No. ED 174 728)

Oakes, J. (1985). Keeping track. New Haven: Yale University Press.

Peterson, P., Wilkinson, L. C., Spinelli, F., & Swing,

S. (1984). Merging the process-product and the sociolinguistic paradigms: Research on small-group processes. In Peterson, P., Wilkinson, L. c., & Hallinan, M. (Eds.), The social context of instruction: Group organization and group processes. Orlando: Academic Press, 3-11.

Rosenbaum, J. E. (1980). Social implications of education grouping. Review of Research in Education, 8, 361-401.

Schafer, W. E. & Olexa, C. (1971). Tracking and opportunity. Scranton, PA: Chandler.

Scheutz, P. (1980). The instructional effectiveness of classroom aides. Pittsburgh, PA: University of Pittsburgh Learning Research and Development Center.

Shuy, R. W. (1978). Toward Future Study of the effect of school setting on learning. In Perspectives on the Instructional Dimensions Study: A supplemental report from the National Institute of Education (pp. 40-48). (ERIC Document Reproduction Service No. ED 174 728)

Slavin, R. E. (1979). Integrating the desegregated classroom: Actions speak louder than words. Educational Leadership, 36, 322-324.

Slavin, R. E. (1980). Cooperative learning. Review of Educational Research, 50(2), 315-342

Slavin, R. E. (1984). Meta-analysis in education: How has it been used? Educational Researcher, 13(8), 6-15.

Slavin, R. E. (1986a). Best evidence synthesis: An alternative to meta-analytic and traditional reviews. Educational Researcher, 15(9), 5-11.

Slavin, R. E. (1986b). Ability grouping and student achievement in elementary schools: A best-evidence synthesis, Report No. 1. Center for Research on Elementary & Middle Schools: The Johns Hopkins University.

Slavin, R. E. & Madden, N. A. (1979). School practices that improve race relations. American Educational Research Journal, 16(2), 169-180.

Smith, M. L. & Glass, G. V. (1980). Meta-analysis of research on class size and its relationship to attitudes and instruction. American Educational Research Journal, 17(4).

Stickney, B. D. & Plunkett, V. R. L. (1982). Has Title I done its job? Educational Leadership, 64, 378-383.

Stonehill, R. M. & Anderson, J. I. (1982). An evaluation of ESEA Title I Program operations and educational effects. Washington, DC: U.S. Department of Education.

Swing, S. & Peterson, P. (1982). The relationship of

student ability and small-group interaction to student achievement. American Educational Research Journal, 19, 259-274.

Tobias, S. (1982). When do instructional methods make a difference? Educational Researcher, 11(4), 4-9.

Vanecko, J. J. & Ames, N. L. with Archambault, F. X. (1980). Who benefits from federal education dollars? Cambridge, MA: Abt Books.

Vasquez-Nuttall Associates, (1983). Chapter 1 elementary reading programs. Interim evaluation report. (ERIC Document Reproduction Service No. ED 238 968)

Wang, M. (1980, April). Evaluating the effectiveness of compensatory education. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Wang, M. M., Hoepfner, R., Zagorski, H., Hemenway, J. A., Brown, D. S., & Bear, M. B. (1978). The nature and recipients of compensatory education. (Technical Report No. 5). Santa Monica, CA: System Development Corporation.

Wargo, M. J., Tallmadge, G. K., Michaels, D. D., Lipe, D., & Morris, S. J. (1972). ESEA Title I: A reanalysis and synthesis of evaluation data from fiscal year 1965 through 1970. Palo Alto, CA:

American Institute for Research.

Webb, N. (1982). Peer interaction and learning in cooperative small groups. Journal of Educational Psychology, 76, 33-44.

Webb, N. (1985). Verbal interaction and learning in peer-directed groups. Theory Into Practice, 24, 32-39.

Webb, N. & Kendarski, C. (1984). Student interaction and learning in small group and whole-class settings. In P. Peterson, L. S. Wilkinson, & M. Hallinan (Eds.), The social context of instruction: Group organization and group processes. Orlando: Academic Press, 153-170.

Wilkinson, L. (1986, June). Grouping students for instruction. Paper presented at the Office of Educational Research and Improvement sponsored and Research and Evaluation Associates, Inc. conducted Conference on the Effects of Alternative Designs in Compensatory Education, Washington, DC.

Yap, K. O. (1983). Effects of instructional setting and approach in compensatory education: A statewide analysis. (ERIC Document Reproduction Service No. ED 230 645)

Yap, K. O. (1985). A cost-analytic approach to determining Chapter 1 program impact: Some

preliminary findings. (ERIC Document Reproduction Service No. ED 254 564)